

AD-A096 412

TEMPLE UNIV PHILADELPHIA PA DEPT OF STATISTICS

F/6 12/1

A GENERAL APPROACH TO LIMITING NORMALITY OF THE PRODUCT-LIMIT E--ETC(U)

FEB 81 A J QUINZI, T M SMITH

AFOSR-78-3673

UNCLASSIFIED

AFOSR-TR-81-0216

NI

| OF |  
AD-A  
096412




END

DATE

FILED

4-81

DTIC

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

## REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS  
BEFORE COMPLETING FORM

1. REPORT NUMBER

AFOSR-TR-81-0216

2. GOVT ACCESSION NO.

AD-A096412

3. REPORTING ORGANIZATION NUMBER

4. TITLE (and Subtitle)

A General Approach to Limiting Normality  
of the Product-Limit Estimator.

5. TYPE OF REPORT &amp; PERIODICITY

Interim

PERFORMING ORG. REPORT NUMBER

7. AUTHOR(S)

A. J. Quinzi and T. M. Smith

8. CONTRACT OR GRANT NUMBER(S)

9. PERFORMING ORGANIZATION NAME AND ADDRESS

Temple University  
Department of Statistics  
Philadelphia, PA 1912210. PROGRAM ELEMENT, PROJECT, TASK  
AREA & WORK UNIT NUMBER(S)

61102F 2304A5

11. CONTROLLING OFFICE NAME AND ADDRESS

Air Force Office of Scientific Research/NM  
Bolling AFB, Washington, DC 20332

12. REPORT DATE

May, 1980

13. NUMBER OF PAGES

15

14. MONITORING AGENCY NAME &amp; ADDRESS (if different from Controlling Office)

11 Feb 81

12 16

15. SECURITY CLASS. (This report)

UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING  
SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

competing risks,  
product-limit estimatorDTIC  
ELECTE

MAR 17 1981

D

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Langberg, Proschan and Quinzi obtain strongly consistent estimators for the unobservable marginal distributions of interest in the competing risks problem. These estimators resemble those of Kaplan and Meier but are appropriate when (a) the risks are dependent and (b) death may result from simultaneous causes. We establish asymptotic normality of these estimators. Our result thereby extends that of Breslow and Crowley from the case of a continuous survival function to an arbitrary survival distribution. This preliminary report represents work currently in progress.

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

81 3 16 023

AD A 096412

FILE COPY

A GENERAL APPROACH TO LIMITING NORMALITY  
OF THE PRODUCT-LIMIT ESTIMATOR

by

A. J. Quinzi<sup>1,2</sup> and T. M. Smith<sup>1</sup>

Preliminary Report

May, 1980

Revised February, 1981

Accession For	
NTIS GRA&I	<input checked="checked" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A	

<sup>1</sup>  
Temple University

<sup>2</sup>  
Research supported by the Air Force Office of Scientific Research  
under AFOSR-~~78-3673~~.

AMS Subject Classification Nos. 60N05, 62E20, 62G05, 62F05, 62P10

Key words: competing risks, product-limit estimator

DTIC  
ELECTE  
MAR 17 1981  
S D

A General Approach to Limiting Normality  
of the Product-Limit Estimator

by

A. J. Quinzi<sup>1,2</sup> and T. M. Smith<sup>1</sup>

ABSTRACT

Langberg, Proschan and Quinzi [Ann. Statist. (1981), to appear] obtain strongly consistent estimators for the unobservable marginal distributions of interest in the competing risks problem. These estimators resemble those of Kaplan and Meier [J. Amer. Statist. Assoc. (1958) 63] but are appropriate when (a) the risks are dependent and (b) death may result from simultaneous causes. We establish asymptotic normality of these estimators. Our result thereby extends that of Breslow and Crowley [Ann. Statist. (1974) 2] from the case of a continuous survival function to an arbitrary survival distribution. This preliminary report represents work currently in progress.

---

<sup>1</sup>  
Temple University

<sup>2</sup>  
Research supported by the Air Force Office of Scientific Research under AFOSR ~~78-3673~~-78-3673A.

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)  
NOTICE OF TRANSMITTAL TO ODC  
This technical report has been reviewed and is  
approved for public release IAW AFR 190-12 (7b).  
Distribution is unlimited.  
A. D. BLUSE  
Technical Information Officer

## 0. Introduction and Summary

In the classical theory of competing risks [cf. the excellent monograph of Birnbaum (1979)] it is assumed that (a) the risks, i.e., the random variables of interest are independent and (b) death does not result from simultaneous causes. The classical estimator for the marginal distributions of interest in the competing risks problem is that of Kaplan and Meier (1958) or generalizations thereof [cf. Peterson (1975, 1977)]. Langberg, Proschan, and Quinzi (1981) [hereafter referred to as LPQ(1981)] obtain strongly consistent estimators for the unobservable marginal distributions of interest when assumptions (a) and (b) above fail to hold. These estimators resemble those of Kaplan and Meier (1958). In Section 1, we examine the competing risks problem in the presence of dependent risks and state a number of known results. In Section 2, we establish the asymptotic normality of the LPQ(1981) estimators. Only the outline of proof is given. This preliminary report represents work currently in progress.

### 1. The competing risks model

Let there be a finite number of causes of death labelled  $1, \dots, r$ . We associate with each cause  $j$  a nonnegative random variable  $T_j$ ,  $j = 1, \dots, r$ . The random variable  $T_j$  represents the age at death if cause  $j$  were the only cause present in the environment. The complete collection of random variables  $T_1, \dots, T_r$  is not observed. Instead, only two quantities are observed: the age at death given by  $\mathcal{T} = \min(T_1, \dots, T_r)$  and the cause of death, labelled  $\xi$ , given by  $I \in \mathcal{I}$  such that  $\xi(\underline{T}) = I$ , where  $\mathcal{I}$  represents the collection of nonempty subsets of  $\{1, \dots, r\}$ . Thus,  $\xi(\underline{T}) = I$  if and only if  $\mathcal{T} = T_i$  for each  $i \in I$  and  $\mathcal{T} \neq T_i$  for each  $i \notin I$ . When death results from exactly one of the  $r$  possible causes, as is usually assumed, then  $\xi$  is the index  $i$  for which  $\mathcal{T} = T_i$ . The biomedical researcher is interested in making inferences about the unobservable random variables  $T_1, \dots, T_r$  by using information from the observable quantities, namely the life length  $\mathcal{T}$  and cause of death  $\xi$ . In particular, he seeks to estimate the  $2^r - 1$  survival probabilities

$$\bar{M}_J(t) = P[\min(T_j, j \in J) > t], J \in \mathcal{I}.$$

We use the following notation throughout. If  $T$  is a nonnegative random variable with distribution function  $F$ , then  $\bar{F} = 1 - F$ .

LPQ(1978) prove the following

Theorem 1.1. Let  $\mathcal{T} = \min(T_1, \dots, T_r)$ , where  $T_1, \dots, T_r$  are nonnegative random variables. Define  $\bar{F}(t, I) = P(\mathcal{T} > t, \xi(\underline{T}) = I)$ ,  $F(t, I) = P(\mathcal{T} \leq t, \xi(\underline{T}) = I)$ ,  $I \in \mathcal{I}$ ,  $\bar{F}(t) = \sum_{I \in \mathcal{I}} \bar{F}(t, I)$  and  $F(t) = 1 - \bar{F}(t)$ .

Then the following statements hold:

(i) A necessary and sufficient condition for the existence of a set of independent random variables  $\{H_I, I \in \mathcal{I}\}$  which satisfy

$$P(\tau > t, \xi(\underline{T}) = I) = P[\min(H_I, I \in \mathcal{I}) > t, H_I < H_J, \text{ each } J \neq I]$$

is that the functions  $F(\cdot, I), I \in \mathcal{I}$ , have no common discontinuities in the interval  $[0, \alpha(F))$ , where  $\alpha(F) = \sup \{t: 1 - F(t) > 0\}$ .

(ii) The random variables  $\{H_I, I \in \mathcal{I}\}$  in (i) have corresponding survival probabilities  $\{\bar{G}_I(\cdot), I \in \mathcal{I}\}$ ,  $\bar{G}_I(t) = P(H_I > t)$ , which are uniquely defined on the interval  $[0, \alpha(F)]$  as follows:

$$(1.1) \quad \bar{G}_I(t) = \prod_{a \leq t} [\bar{F}(a)/\bar{F}(a^-)] \cdot \exp \left[ - \int_0^t dF^C(\cdot, I)/\bar{F} \right], \quad 0 \leq t \leq \alpha(F),$$

where  $F^C(\cdot, I)$  is the continuous part of  $F(\cdot, I)$ , the product is over the discontinuities  $\{a\}$  of  $F(\cdot, I), I \in \mathcal{I}$ , and the product over an empty set is defined as unity.

Remark 1.2. Although motivated by the competing risks model, Theorem 1.1 applies to any model where observations include (1) the time at which a particular event occurs and (2) the identity of the causes (among a finite number) which result in the occurrence of the event. For example, suppose a personnel study is undertaken to study the departure patterns of employees in a large company. The data on each employee might consist of (1) length of stay, i.e., the time from arrival to termination, and (2) the reason for termination. Here, each employee terminates (dies) for one or more of several reasons (causes).

Remark 1.3. Formula (1.1) represents each distribution in the independent collection  $\{H_I, I \in \mathcal{I}\}$  as a function of the (observable) cause-specific subdistribution functions  $F(\cdot, I)$ ,  $I \in \mathcal{I}$ , as well as the (observable) survival function  $\bar{F}(t) = P(\mathcal{T} > t)$ . It is this representation of distributions in the independent collection by observable functions which plays a key role in the estimation problem.

Let  $\underline{T}_i = (T_{1i}, \dots, T_{ri})$ ,  $i = 1, \dots, n$ , represent a random sample from the joint distribution of the nonnegative random variables  $T_1, \dots, T_r$ . For each  $J \in \mathcal{I}$ , let  $\bar{M}_J(t) = P[\min(T_j, j \in J) > t]$ . For  $j \in \{1, \dots, r\}$ , we write  $M_j(t)$  instead of  $M_{\{j\}}(t)$ . For each  $i = 1, \dots, n$ , only  $\mathcal{T}_i$  and  $\xi_i$  are observed, where  $\mathcal{T}_i = \min(T_{1i}, \dots, T_{ri})$  and  $\xi_i = J$  whenever  $\mathcal{T}_i = T_{ji}$  for each  $j \in J$  and  $\mathcal{T}_i \neq T_{ji}$  for each  $j \notin J$ . It is important

to note that we have not made either of the two classical assumptions, namely (a) the risks, i.e., the random variables  $T_1, \dots, T_r$  are independent; and (b) death does not result from simultaneous causes, i.e.,  $P(T_i = T_j) = 0$  for  $i \neq j$ . If (a) and (b) hold, the function  $\bar{M}_J(t)$  may be estimated (consistently) [cf. Peterson (1975)] by using a generalized version of the Kaplan-Meier (1958) (product-limit) estimator

$$(1.2) \quad \hat{\bar{M}}_J(t) = \prod_i [(n - i)/(n - i + 1)],$$

where the product is over the ranks  $i$  of those ordered observations  $\mathcal{T}_{(i)}$  such that  $\mathcal{T}_{(i)} \leq t \leq \mathcal{T}_{(n)}$  and  $\mathcal{T}_{(i)}$  corresponds to a death from at



least one cause  $j \in J$ . If  $\tau_{(n)}$  corresponds to a death from a cause  $j \in J$ , then  $\hat{M}_j(t)$  is defined to be zero for  $t > \tau_{(n)}$ . Otherwise,  $\hat{M}_j(t)$  is undefined for  $t > \tau_{(n)}$ . [In the original formulation by Kaplan and Meier (1958),  $r = 2$  and  $T_1$  corresponded to the time until death, while  $T_2$  corresponded to the time at which a loss occurred.]

Suppose now that it is not assumed that  $T_1, \dots, T_r$  are independent. LPQ(1981) prove the following

Theorem 1.4. Let  $T_1, \dots, T_r$  be nonnegative random variables such that the functions  $F(t, I) = P(\tau \leq t, \xi(\tau) = I)$ ,  $I \in \mathcal{I}$ , have no common discontinuities. Define  $\mathcal{I}_I = \{J \in \mathcal{I} : J \cap I \neq \emptyset\}$ . Fix  $I \in \mathcal{I}$ . Then for each  $t \in [0, \alpha(F)]$ ,

$$(1.3) \quad \bar{M}_I(t) = \prod_{J \in \mathcal{I}_I} \bar{G}_J(t)$$

if and only if

$$(1.4a) \quad \bar{M}_I(a)/\bar{M}_I(a^-) = \begin{cases} \bar{F}(a)/\bar{F}(a^-), & a \in D(\mathcal{I}_I) \\ 1, & \text{otherwise;} \end{cases}$$

and

$$(1.4b) \quad P(\tau_{I'} > t \mid \tau_I = t) = P(\tau_{I'} > t \mid \tau_I > t),$$

where  $\bar{G}_J$  is given by (1.1),  $D(\mathcal{I}_I)$  is the set of discontinuities of the function  $F(t, \mathcal{I}_I) = P(\tau \leq t, \xi(\tau) \in \mathcal{I}_I)$ ,  $\tau_I = \min(T_i, i \in I)$  and  $I'$  is the complement of  $I$  in  $\{1, \dots, r\}$ .

Remark 1.5. By finding consistent estimators for the functions  $\bar{G}_j$  in (1.3), LPQ(1981) establish that (1.4a) and (1.4b) are necessary and sufficient conditions on the joint distribution of  $T_1, \dots, T_r$  for the existence of a consistent estimator of  $\bar{M}_I$  in (1.3).

Remark 1.6. Suppose that the random variables  $T_I = \min(T_i, i \in I), I \in \mathcal{I}$ , have absolutely continuous distributions. Let  $m_I(t)$  [respectively,  $\bar{M}_I(t)$ ] and  $m_{I|I'}(t)$  [respectively,  $\bar{M}_{I|I'}(t)$ ] denote the density (respectively, survival) function and conditional density (respectively, conditional survival) function of  $T_I$  and  $T_{I'}$  given  $T_{I'} > t$ . Then condition (1.4b) is equivalent to

$$m_{I|I'}(t)/\bar{M}_{I|I'}(t) = m_I(t)/\bar{M}_I(t).$$

In other words, the conditional failure rate function of  $T_I$  given  $T_{I'} > t$  is equal to the (unconditional) failure rate function of  $T_I$ . Stated differently, the random variables  $T_I$  and  $T_{I'}$  are independent "along the diagonal  $T_I = T_{I'}$ ". This property of "diagonal independence" is of importance in the case of dependent competing risks and is presently being studied by the authors. Desu and Narula (1977) arrive at a condition similar to (1.4b) in the special case when  $T_1, \dots, T_r$  have a joint distribution which is absolutely continuous.

Suppose now that the functions  $F(t, I), I \in \mathcal{I}$ , have no common discontinuities. We make no assumption as to the independence of  $T_1, \dots, T_r$ . In view of (1.3), a natural estimator for  $\bar{M}_I$  is

$$(1.5) \quad \tilde{\bar{M}}_I(t) = \prod_{j \in \mathcal{I}_I} \bar{G}_{j,n}(t).$$

where  $\bar{G}_{J,n}(t)$  is obtained from the right side of (1.1) by replacing

$F(\cdot, I)$  and  $F$  by their empirical counterparts

$$F_n(t, I) = n^{-1} \sum_{i=1}^n \chi_{\{\tau_i \leq t, \xi_i = I\}}$$

and

$$F_n(t) = n^{-1} \sum_{i=1}^n \chi_{\{\tau_i \leq t\}},$$

where  $\chi_A$  is the indicator function of the set  $A$ . LPQ(1981) show that in

this case, (1.5) is a (strongly) consistent estimator for  $\bar{M}_T$  when (1.4a,b) hold.

Remark 1.7. If  $T_1, \dots, T_r$  are independent and  $P(T_i = T_j) = 0$  for  $i \neq j$ , then (1.5) reduces to the usual Kaplan-Meier (1958) estimator (1.2) or a version thereof.

Remark 1.8. Suppose for a moment that we make no assumption on the underlying distribution of  $T_1, \dots, T_r$  except that the functions  $F(t, I)$ ,  $I \in \mathcal{I}$ , have no common discontinuities. Let  $0 \equiv \tau_{(0)} \leq \dots \leq \tau_{(n)} < \tau_{(n+1)} \equiv \infty$

denote the ordered values of times  $\tau_1, \dots, \tau_n$  at which deaths occur.

We do not exclude the possibility of multiple deaths at  $\tau_{(j)}$ . We thus obtain the (possibly degenerate) intervals  $[0, \tau_{(1)}), [\tau_{(1)}, \tau_{(2)}), \dots,$

$[\tau_{(n)}, \infty)$  such that the number of deaths in any interval is exactly one.

For each interval  $[\tau_{(j)}, \tau_{(j+1)})$ , estimate the proportion  $p_j$  of individuals alive just after  $\tau_{(j)}$  that survive the interval as follows:

let  $N(t)$  = the number of individuals observed and surviving  
 at  $t$ , when deaths due to cause  $I$  (but not deaths due  
 to any other cause) at  $t$  itself are subtracted off;  
 and  $\delta_j = N(\tau_{(j)}^-) - N(\tau_{(j)}) =$  the number of deaths at  $\tau_{(j)}$ .

Then the estimate of  $p_j$  above is

$$\hat{p}_j = [N(\tau_{(j)}^-) - \delta_j] / N(\tau_{(j)}^-).$$

Now, to estimate the probability of surviving until  $t$  if cause  $I$  were  
 the only risk present in the environment, Kaplan and Meier (1958) calculate

$$(1.6) \quad \bar{M}_I^*(t) = \prod_{j=1}^k \hat{p}_j.$$

For any given set of data, formula (1.6) yields the same numerical estimate as formula (1.5). Recall, however, that (1.5) is a consistent estimator of  $\bar{M}_I$  if and only if (1.4a,b) hold. Yet, even in the face of ignorance about the truth or falsity of (1.4a,b), we know precisely what parameter of the underlying distribution is being estimated (consistently) by (1.6), namely

$$(1.7) \quad \prod_{j \in \mathcal{J}_I} \bar{G}_j(t),$$

where  $\bar{G}_j$  is given by (1.1). To the authors' knowledge, this fact has never been pointed out.

## 2. The main result

In this section we outline a proof of the fact that, viewed as a process in  $t$ , the estimator (1.5) converges to a Gaussian process. As a result, we extend a result of Breslow and Crowley (1974) from the case of a continuous survival to an arbitrary survival function. For simplicity, we assume here that the distribution  $F$  of  $\mathcal{T}$  has finitely many discontinuities. The case of a countable infinity of discontinuities will be presented in a subsequent report.

We inquire into the asymptotic distribution of

$$\widetilde{M}_I(t) = \prod_{J \in \mathcal{J}_I} \bar{G}_{J,n}(t) = \prod_{J \in \mathcal{J}_I} \prod_{a \leq t} [\bar{F}_n(a)/\bar{F}_n(a^-)],$$

where the last product is over the set of observations  $\{a\}$  such that  $\mathcal{T}_i = a$  and  $\mathcal{J}_i = J$ ,  $i = 1, \dots, n$ . Let this set  $\{a\}$  of points be denoted by  $D(n, J)$  and let  $C(J)$  [ $D(J)$ ] be the set of continuities (discontinuities) of the function  $F(t, J)$ . Then we can write

$$\begin{aligned} (2.1) \quad \sqrt{n}[\bar{G}_{n,J}(t) - \bar{G}_J(t)] &= \sqrt{n}[e^{H_{n,J}(t)} - e^{H_J(t)}] \\ &= \sqrt{n}[H_{n,J}(t) - H_J(t)] e^{H(t)} + \frac{\sqrt{n}}{2}[H_{n,J}(t) - H_J(t)]^2 e^{H^*(t)}, \end{aligned}$$

where  $H_{n,J}(t) = \sum_{a \leq t} \ln [\bar{F}_n(a)/\bar{F}_n(a^-)] \chi_{D(n, J)}(a)$ .

$H_J(t) = \sum_{a \leq t} \ln [\bar{F}(a)/\bar{F}(a^-)] \chi_{D(J)}(a) - \int_0^t dF^C(\cdot, J)/\bar{F}$ , and the function  $H^*(t)$  is between  $H_{n,J}(t)$  and  $H_J(t)$ .

We consider first the asymptotic distribution of  $\sqrt{n}[H_{n,J}(t) - H_J(t)]$ .

We have

$$\sqrt{n}[H_{n,J}(t) - H_J(t)] = \sqrt{n}[A_{n,J}(t) - A_J(t)] + \sqrt{n}[B_{n,J}(t) - B_J(t)],$$

$$\text{where } A_{n,J}(t) = \sum_{a \leq t} \ln [\bar{F}_n(a)/\bar{F}_n(a^-)] \chi_{D(n,J)}^{(a)} \cdot \chi_{D(J)}^{(a)},$$

$$A_J(t) = \sum_{a \leq t} \ln [\bar{F}(a)/\bar{F}(a^-)] \chi_{D(J)}^{(a)},$$

$$B_{n,J}(t) = \sum_{a \leq t} \ln [\bar{F}_n(a)/\bar{F}_n(a^-)] \chi_{D(n,J)}^{(a)} \cdot \chi_{C(J)}^{(a)}, \text{ and}$$

$$B_J(t) = - \int_0^t dF^C(\cdot, J)/\bar{F}.$$

We can now state

**Theorem 2.1.** Assume that each function  $F(\cdot, J)$ ,  $J \in \mathcal{J}$ , has finitely many discontinuities. Fix  $J \in \mathcal{J}$  and let  $0 < a_1 < \dots < a_k < \infty$  denote the discontinuities of  $F(\cdot, J)$ . Then the  $k$ -dimensional random vector whose  $i$ th component is

$$\sqrt{n} \sum_{j=1}^i \left\{ \ln [\bar{F}_n(a_j)/\bar{F}_n(a_j^-)] - \ln [\bar{F}(a_j)/\bar{F}(a_j^-)] \right\}$$

converges in distribution to a  $k$ -dimensional multivariate normal with mean vector  $\underline{0}$  and covariance matrix  $\Sigma = (\sigma_{ij})$ , which can be represented thus:

$$\Sigma = \begin{bmatrix} b_1 & b_1 & b_1 & \dots & b_1 \\ b_1 & b_1+b_2 & b_1+b_2 & \dots & b_1+b_2 \\ b_1 & b_1+b_2 & b_1+b_2+b_3 & \dots & b_1+b_2+b_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_1 & b_1+b_2 & b_1+b_2+b_3 & \dots & b_1 + \dots + b_k \end{bmatrix},$$

where  $b_i = [F(a_i)/\bar{F}(a_i)] - [F(a_i^-)/\bar{F}(a_i^-)]$ ,  $i = 1, \dots, k$ .

The proof of Theorem 2.1 is straightforward and is omitted.

Now define a process  $Z_{J,1}(t)$  in  $D = D[0, \alpha(F)]$  whose finite-dimensional distributions are multivariate normal with  $EZ_{J,1}(t) = 0$  and

$$\text{Cov}[Z_{J,1}(s), Z_{J,1}(t)] = \begin{cases} \sigma_{ij} & \text{for } s \in [a_i, a_{i+1}), t \in [a_j, a_{j+1}), s \leq t \\ \sigma_{ji} & \text{for } s > t; \end{cases}$$

where  $\sigma_{ij}$  is the  $(i,j)$ th entry of  $\Sigma$  in Theorem 2.1 and  $D[0, \alpha(F)]$  is

the space of functions on  $[0, \alpha(F)]$  that are right-continuous and have left-hand limits. Such a process exists by Theorem 15.3 of Billingsley (1968).

**Theorem 2.2.** The process  $\sqrt{n}[A_{n,J}(t) - A_J(t)]$  converges weakly to  $Z_{J,1}(t)$  as  $n \rightarrow \infty$ .

Proof. Note that  $\sqrt{n}[A_{n,J}(t) - A_J(t)]$

$$= \sqrt{n} \sum_{i=1}^J \left\{ \ln [\bar{F}_n(a_i)/\bar{F}_n(a_i^-)] X_{D(J,n)}(a_i) - \ln \bar{F}(a_i)/\bar{F}(a_i^-) \right\}$$

for  $t \in [a_j, a_{j+1})$ . For  $\delta \leq \min_{1 \leq i,j \leq k} |a_i - a_j|$ , it is easily seen

that  $w_x^n(\delta) = \sup \min (|x_n(t) - x_n(t_1)|, |x_n(t_2) - x_n(t)|) = 0$ ,

where  $x_n(t) = \sqrt{n}[A_{n,J}(t) - A_J(t)]$  and the supremum extends over  $t, t_1, t_2$

such that  $t_1 \leq t \leq t_2$  and  $t_2 - t_1 \leq \delta$ . The theorem follows from

Theorem 2.1 above and Theorem 15.4 of Billingsley (1968). ||

Breslow and Crowley (1974) show that the pair  $(X_n, Y_n) \in D[0, \alpha(F)] \times D[0, \alpha(F)]$  defined by  $X_n = \sqrt{n}(\bar{F}_n - \bar{F})$ ,  $Y_n = \sqrt{n} \left[ \sum_{a \leq t} X_{C(J)}(a) X_{D(J,n)}(a)/n - F^G(\cdot, J) \right]$  converges weakly to a bivariate Gaussian process  $(X, Y)$  which has mean vector zero and a covariance structure given by

$$(2.2) \quad \begin{cases} \text{Cov}(X(s), X(t)) = F(s)\bar{F}(t), \text{Cov}(Y(s), Y(t)) = F^G(s, J) [1 - F^G(t, J)] \\ \text{Cov}(Y(s), X(t)) = F^G(s, J)\bar{F}(t), \text{Cov}(X(s), Y(t)) = F^G(s, J) - F(s)F(t, J). \end{cases}$$

Thus, by Theorem 4 of Breslow and Crowley (1974), the process

$\sqrt{n}[B_{n,J}(t) - B_J(t)]$  converges weakly to the Gaussian process  $Z_{J,2}(t)$  defined by

$$Z_{J,2}(t) = \int_0^t [X/\bar{F}^2] dF^C(\cdot, J) + [Y(t)/\bar{F}(t)] - \int_0^t Y d(1/\bar{F}),$$

where  $(X, Y)$  is the bivariate mean 0 Gaussian process satisfying (2.2).

Furthermore, the covariance structure of the limiting process  $Z_{J,2}(t)$  can be obtained in a manner similar to that in Breslow and Crowley (1974).

Combining this result with Theorem 2.2 above, we have

Theorem 2.3. The process  $\sqrt{n}[H_{n,J}(t) - H_J(t)]$  converges weakly to the Gaussian process  $Z_{J,1}(t) + Z_{J,2}(t) \equiv Z_J(t)$ .

Remark 2.4. The covariance structure of the limiting process in Theorem 2.3, as well as in the remaining theorems, may be obtained in a tedious but straightforward manner. The exact derivations are given in a later report.

Consider now (2.1). Since  $\sqrt{n}[H_{n,J}(t) - H_J(t)]$  converges weakly, the second term in (2.1) converges to 0 in probability. Thus, we have

Theorem 2.5. The process  $\sqrt{n}[G_{n,J}(t) - G_J(t)]$  converges weakly to the Gaussian process  $Z_J(t) \cdot \bar{G}_J(t)$ , where  $Z_J(t)$  is the limiting process in Theorem 2.3.

Remark 2.6. Finally, by an application of the so-called  $\delta$ -method [cf. Rao (1973)], we see that the estimator  $\tilde{M}_J(t)$  given by (2.2) also converges weakly to a Gaussian process.



# REFERENCES

- [1] Billingsley, P. (1968). Convergence of Probability Measures. John Wiley, New York.
- [2] Birnbaum, Z. W. (1979). On the mathematics of competing risks. DHEW Publication No. (PHS) 79-1351, Vital and Health Statistics Data Evaluation and Methods Research, Series 2, No. 77.
- [3] Breslow, N. and Crowley, J. (1974). A large sample study of the life table and product-limit estimates under random censorship. Ann. Statist. 2, 437-453.
- [4] Desu, N. M. and Narula, S. C. (1977). Reliability estimation under competing causes of failure. The Theory and Applications of Reliability I, ed. I. Shimi and C. P. Tsokos, Academic Press, New York.
- [5] Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. J. Amer. Statist. Assoc. 53, 457-481.
- [6] Langberg, N., Proschan, F. and Quinzi, A. J. (1978). Converting dependent models into independent ones, preserving essential features. Ann. Prob. 6, 174-181.
- [7] Langberg, N., Proschan, F. and Quinzi, A. J. (1981). To appear in Ann. Statist.
- [8] Peterson, A. V., Jr. (1975). Nonparametric estimation in the competing risks problem. Stanford University Technical Report No. 13.
- [9] Peterson, A. V., Jr. (1977). Expressing the Kaplan-Meier estimator as a function of empirical subsurvival functions. J. Amer. Statist. Assoc. 72, 854-858.
- [10] Rao, C. R. (1973). Linear Statistical Inference and Its Applications. John Wiley, New York.

ATE  
LMED  
-8